INJURY SURVIVABILITY

Peter G. Martin, Rolf H. Eppinger
National Highway Traffic Safety Administration
United States of America
Paper Number 306

ABSTRACT

A procedure is presented to estimate the risk to life that multiple injuries pose to crash victims. It continues Eppinger's original work first presented at the 1981 ESV conference. At its core, the procedure uses only the two most serious injuries – denoted as the primary injury and the secondary injury – to characterize a victim's entire injury record. Nine years of data from the Crashworthiness Data System (CDS) containing injury records for over 50,000 crash victims – including over 3500 fatalities – are analyzed. For each victim, the top two injuries are defined by using a data-driven approach based on actual CDS outcomes as opposed to relying solely on the Abbreviated Injury Scale, a heuristic ranking system developed by a panel of experts. Results show that for a given primary injury, the risk to life varies profoundly depending upon the secondary injury. Victim age has a substantial effect, too. When deviance statistics are considered, the new procedure predicts fatalities better than other injury scales (including the Injury Severity Score). Ultimately, this two-injury procedure promotes better estimates of safety benefits by directly quantifying and specifying fatality-related injuries in the CDS data.

INTRODUCTION

Research Philosophy at NHTSA. The National Highway Traffic Safety Administration (NHTSA) is responsible for reducing deaths, injuries, and economic losses resulting from motor vehicle crashes. This is accomplished by setting and enforcing safety performance standards for motor vehicles and through grants to state and local governments to enable them to conduct effective local highway safety programs. Within NHTSA's Advanced Safety Research division, these aims are achieved by taking a "data driven" approach in research activities that will lead to a reduction in crashes and their consequences. As such, decisions are grounded in sound statistical and engineering methods. Generally, there must be enough existing data to show that a proposed countermeasure will reduce the risk of injuries significantly before decisions are made and changes implemented. To aid in such assessments, NHTSA maintains epidemiological data on the nature, causes, and injury outcomes of crashes.

The National Automotive Sampling System - Crashworthiness Data System. The Crashworthiness Data System (CDS) is one of the epidemiological databases maintained by NHTSA. The CDS is a nationally representative probability sample of police-reported automobile crashes in the United States. CDS cases are limited to crashes that involve at least one passenger car that was towed from the crash scene due to damage resulting from the crash. Each year, the CDS collects data on about 5000 crashes from 24 geographic sites across the United States. CDS case files are assembled by crash investigators by referring to police reports and hospital records. Each case is assigned a weighting factor that represents an estimate of the number of like-mannered cases that occurred during the sample year. Investigators also conduct crash victim interviews, visit the crash site, and inspect the post-crash vehicles. Over 300 coded CDS variables describe the occupants, injuries, and vehicles involved in the crash.

Abbreviated Injury Scale. The CDS is particularly useful to NHTSA researchers in establishing priorities for the development of crash test dummies. Since 1993, a six-digit code has been assigned to each occupant injury in accordance with the CDS Injury Coding Manual (Benton, 1993). This code – which may be cross-referenced with detailed nomenclature in the coding manual – defines injury specifics such as the body region, organ, and type of lesion. Furthermore, each code is appended with an injury severity suffix in accordance with the 1990 revision of the Abbreviated Injury Scale (AIS-90) developed under the auspices of the Association for the Advancement of Automotive Medicine (AAAM, 1990). This suffix (commonly referred to as the AIS level) takes on a numerical value ranging from 1 to 6 corresponding to the injury severity: 1=minor, 2=moderate, 3=serious, 4=severe, 5=critical, 6=maximum severity. (If a motorist suffers an injury of an unknown type, a special code is assigned with a suffix of 7, which indicates an injury of unknown severity.)

By examining the severity and frequency of various types of injuries, researchers may establish performance test and dummy instrumentation requirements. CDS injury data are used, for example, to justify the enactment of the new performance requirements for child restraint systems in frontal crashes (NHTSA, 2002). The data are used to estimate how new dynamic performance requirements – based on a variety of new child dummy measurands – will prevent injuries and save lives.

Injury Characterization Tactics. In establishing research priorities, a thorough injury assessment must be carried out to determine how a proposed dummy requirement will lead to a reduction in injuries. A typical assessment uses CDS data to show that a proposed performance requirement or countermeasure will result in a significant reduction of injuries. Such an assessment usually requires a more in-depth analysis than simply counting CDS injuries. Instead, crash victims are classified by the single injury that posed the greatest threat-to-life which is defined as the maximum AIS injury, or MAIS injury. When analyzing crash victims in aggregate, the MAIS injury is reasoned to be solely responsible for the victim's impairment, death, or incurred costs. Therefore, MAIS incidence levels reflect the number of people, not the number of injuries. This convention provides an easy, straight-forward way to track population counts while classifying injuries.

The MAIS ranking alone only describes the severity of the maximum injury, not the type of injury it is. Therefore, injury assessments commonly identify the body region to which the MAIS injury corresponds. Thus, a driver who sustained several injuries of AIS severity level 3 (i.e., several AIS 3 injuries) and one AIS 4 brain injury is treated as having had an MAIS 4 brain injury only. If a victim had two injuries with the same AIS value, the victim is characterized by only the most severe injury, which is selected according some body region hierarchy. For example, NHTSA uses a forty-one level hierarchy to select body regions to assess the economic costs of injuries (Blincoe et al, 2002). It ranks many non-life-threatening skeletal fractures relatively high due to associated societal costs such as long-term disability coverages.

The one-injury victim characterization tactic has its limitations. There are sometimes over twenty injuries listed for each CDS occupant, with many injuries over multiple body regions having the same maximum AIS level. Thus, the single, most life-threatening injury is not always apparent. Moreover, the MAIS body regions associated with fatalities tend to be the head and thorax, and rarely the extremities or the abdomen. But abdominal and extremity injuries do, in fact, increase fatality risk. For example, in a blunt traumatic brain injury, the blood replacement requirement of an additional extremity injury has been shown to increase the risk of death (Siegel et al, 1991). Under an MAIS injury assessment, the risk to life of such additional injuries is masked.

Implicit in an MAIS assessment is an assumption of threat-to-life equivalence of like-ranking injuries across body regions. That is, all crash victims characterized by MAIS 4 are assumed to have an identical threat-to-life, and that threat is assumed to be higher than the threat-to-life of any victim characterized by MAIS 3. Eppinger (1987) has found discrepancies in this assumption when specific body regions and fatality incidence levels in actual CDS data are considered. The threat-to-life posed by lesions of equal AIS rank, but residing in different body regions, is not the same. While ranks within a body region are more or less consistent (AIS 5's pose a greater threat than 4's; 4's are riskier than 3's, etc.), ranks across body regions are not. For example, head injuries of AIS 3 pose about the same threat-to-life as abdominal injuries of AIS 4.

CHARACTERIZATION METHOD

Goals and Objectives. The objective of the present study is to characterize the injuries of CDS victims in a way that correlates well with survivability while helping to guide biomechanical research. The injury characterization procedure makes use of a predictive model that discriminates among several types of injuries and body regions. The model is based on the mortality outcomes of actual CDS cases analyzed as a whole. As such, the characterization model lends itself to studying injuries in aggregate, not case-by-case assessments.

General Description. At its core, the new procedure uses only the two most serious injuries – the primary injury and secondary injury – to characterize a victim's entire injury record. Thus, instead of using just a single MAIS injury, the new "Primary/Secondary" model uses two injuries. Whereas the primary injury sets the upper limit of the survival probability, the secondary injury can be thought of as a "survivability modulator". This two-injury approach expands upon the original efforts of Eppinger and Partyka (1981). As in the original analysis, the present approach sorts each CDS case by the two most life-threatening injuries, and stratifies them by the anatomical location. The present approach, however, uses actual CDS outcomes to help select and sort injuries, and makes use of a logistic regression model with separate parameter estimates for the primary and secondary injuries. Moreover, no assumptions are made in the present analysis about threat-to-life equivalence of injuries across body regions.

Sorting and Grouping is the Key. Since just two injuries are chosen to represent each crash victim's entire injury profile, the selection process is crucial to obtain a good injury survivability correlation. First, a ranking system is established whereby all of the six-digit injury descriptor codes (over one thousand codes) are ordered by survivability. In establishing the ranking system, the entire data set is analyzed in aggregate to objectively discriminate among injuries across body regions having the same AIS value.

Secondly, the thousand-plus codes are placed into injury groups with similar survivability rankings. Grouping of injuries is necessary to gain statistical significance. For example, even though the injury described by injury code 120402.5 (basilar artery injury) may be of particular interest, it occurs very rarely – not enough to warrant its own group. On the other hand, code 890402.1 (lower extremity contusion) appears in numerous injury records (enough to warrant its own group), but it is very low on any threat-to-life scale and is not of much interest. Therefore, it is lumped together with other similar minor injuries.

Injury groups, however, should not be too broad. When specific injuries are placed into a broader injury group, injuries having distinct survivability risks will be lumped together under a single banner with an approximated survivability risk for the whole group. Thus, some precision will be lost due to the lumping. If the CDS contained enough cases, then no lumping would be necessary and each specific injury (of the 1000+ injuries) could be treated on its own.

Injury groups should be defined in a way that will help justify biomechanical priorities. As such, each group should contain injuries to the same general body region and they should be related to a particular dummy metric. For example, injuries to the cervical spine and neck may be lumped together (corresponding with upper and lower neck load cells), but cervical and lumber injuries probably should not, because they are not related to the same dummy instrumentation.

Case-by-Case Labeling. Finally, a case-by-case assessment is undertaken in which the multi-injury record of each CDS victim is queried and the top two injuries – and injury groups – are identified based on a previously defined primary and secondary injury sorting and grouping scheme. Thus, victims are characterized as having only two injuries: a primary injury and a secondary injury. It is further hypothesized that a mathematical expression (in the form of a logistic regression model) describes how the hazards of primary and secondary injuries combine to influence the overall survivability of a crash victim. That is,

$$\begin{aligned}\text{Survival Rate} &= 1/(1 + \text{Exp}[-\beta_o - \beta_i{*}P_i - \beta_j{*}S_j]) \\ &= \text{Logit}[\beta_o + \beta_i{*}P_i + \beta_j{*}S_j] \qquad [1]\end{aligned}$$

where:

P_i is the primary injury group (n primary groups)
S_j is the secondary group (m secondary groups)
β_o is the model intercept
β_i is the regression parameter associated with P_i
β_j is the regression parameter associated with S_j

The remainder of this paper demonstrates how a Primary/Secondary model may be constructed for use in a general survivability analysis. Its predictability is then compared with other known survivability indices. The model is used to identify the types of injuries that contribute most to fatalities with an example of how reducing a specific type of injury translates into lives saved. Possibilities for model improvements are also discussed with insights into how the Primary/Secondary scheme may be adapted to help evaluate the benefits of a particular countermeasure.

DATA SET: NASS – CDS 1993-2001

Data Set Overview. The characterization method is based on a data set extracted from nine years of CDS data, 1993-2001. The working data set contains only victims with MAIS 3+ injuries. This serves a dual purpose. Most CDS crash victims suffer only minor, non-life-threatening injuries. By disregarding these victims, a better statistical correlation may be realized for those injuries that are truly life-threatening. It also eliminates rare but highly confounding "undercoded" fatalities. In such a case, the injury record for a fatality is incomplete because there has been no thorough medical examination. (This occurs often when a victim is "dead on arrival".) Therefore, the highest-ranking injury documented on the injury record form is only an AIS 1 or 2. Thus, the injury record is "undercoded". (On the other hand, it is extremely rare for a fatal case to have a fully completed injury record with the highest ranked injury denoted as an AIS 2.)

MAIS 6 cases are also excluded from the data set because they are deterministic; with a few rare exceptions (less than 1%) they are all fatalities. Moreover, only adults (ages 15 and over) are considered since the mortality rates of many types of injuries are known to vary significantly if they occur in children (Sartorelli et al, 1999). In all, the working data set contains data on about 7000 crash victims – including records for more than 1900 fatalities – over the nine-year span. When these figures are weighted to represent national totals, there are about 600,000 people and 100,000 fatalities over the nine years.

PRIMARY/SECONDARY MODEL

A general application of the Primary/Secondary injury characterization method is demonstrated presently. Injury groups are chosen to represent five general anatomical regions: the chest, head, abdomen, lower extremities above the knee, and any other region (upper extremities and below knees). Each region is further classified according to threat-to-life, resulting in a total of ten injury groups as shown in Table 1. The specific injuries encompassed by each group are also denoted in Table 1. The number of injury groups used to represent an anatomical region is based on the relative threat-to-life and incidence levels of injuries within the region. As such, the head and chest – with high threat-to-life variation among many injuries – are represented by three groups each (the subscripts H, M, and L indicate relative threat-to-life – high, moderate, low). The abdomen is represented by two groups and the lower extremities by just one. (Note that these injury groups are selected to carry out a general analysis that is not directed at studying any particular injury or countermeasure. Sometimes a less general assessment is desired with injury groups that differ from the ten listed in Table 1. This is discussed later.)

Table 1. Injury groups and group rankings by threat-to-life.

Rank	Injury Group	Injuries encompassed by the group according to body region and AIS Severity
1	$Chest_H$	Thorax and Thoracic Spine - AIS 5
2	$Head_H$	Head, Brain, Cervical Spine, Neck, Face - AIS 5
3	$Chest_M$	Thorax and Thoracic Spine - AIS 4
4	$Abdo_H$	Abdominal and Abdominal Spine - AIS 4, 5
5	$Head_M$	Head, Brain, Cervical Spine, Neck, Face - AIS 4
6	$Head_L$	Head, Brain, Cervical Spine, Neck, Face - AIS 3
7	$Chest_L$	Thorax and Thoracic Spine - AIS 3
8	$Abdo_L$	Abdominal and Abdominal Spine - AIS 3
9	LowEx	Pelvis, Femur, Leg Amputation - AIS 3, 4, 5
10	Other	All other injuries, or no injury (secondary only)

Once the injury groups are formed, their ranks are determined via a multivariate logistic regression analysis in which actual CDS outcomes are used. Ten new variables representing the ten groups are used to describe the injuries to each CDS crash victim. The new variables take on values of 1 (an injury belonging to the group is present) or 0 (the injury is not present). Using mortality as the dependent variable (1=fatal, 0=non-fatal), a logistic regression analysis is carried out to determine the relative threat-to-life (i.e., the rank) of each injury group. The ranks are listed in Table 1 by descending threat-to-life ("$Chest_H$" injuries have the highest threat-to-life; "Oth" injuries have the lowest).

The injury group ranks are subsequently used to select the primary and secondary injuries of each crash victim. (Note that the working data set has been extracted so that "Oth" injuries are never primary injuries.) Then, two new categorical variables are defined for each victim: a primary injury variable falling into one of nine categories, and a secondary injury variable falling into one of ten categories. Another logistic regression analysis is carried out in accordance with the model of Eq. [1] to determine the estimates of the nineteen corresponding parameters.

RESULTS

<u>Logistic Regression</u>. Table 2 provides a glimpse of the Primary/Secondary groups and their logistic regression parameter estimates. The parameter estimates and their standard errors are found by taking into account the CDS national expansion case weights and applying the SAS logistic regression procedure (SAS, 1999). All of the parameter estimates have a high level of confidence associated with them (i.e., parameter estimate/standard error > 2) as computed by SAS. (Note: SAS software does not account for CDS's multi-stage sampling system in computing standard errors. Properly computed standard errors may lower the level of confidence associated with the parameter estimates.)

Table 2. Parameter estimates for primary and secondary injuries.

Primary Injury			Secondary Injury		
Param.	Injury Group	Estimate	Param.	Injury Group	Estimate
β_1	$Chest_H$	-2.745	β_{10}	$Chest_H$	-0.966
β_2	$Chest_M$	-0.679	β_{11}	$Chest_M$	-0.836
β_3	$Chest_L$	0.497	β_{12}	$Chest_L$	0.100
β_4	$Head_H$	-0.802	β_{13}	$Head_H$	-0.649
β_5	$Head_M$	-0.112	β_{14}	$Head_M$	-0.390
β_6	$Head_L$	0.101	β_{15}	$Head_L$	0.343
β_7	$Abdo_H$	-0.123	β_{16}	$Abdo_H$	-0.072
β_8	$Abdo_L$	2.970	β_{17}	$Abdo_L$	0.872
β_9	LowEx	1.088	β_{18}	LowEx	0.761
β_0	Intercept	1.161	β_{19}	Oth	1.035

The parameters offer a direct measure of survivability: lower parameters indicate greater threat-to-life ($Chest_H$ injuries are the most life-threatening). The overall survivability of an injury group is a function of the sum of a primary parameter estimate and a secondary estimate (see Eq. 1). From Table 2, one can see how the sum – and the overall survivability – of a group containing a given primary injury can vary greatly depending upon the type of secondary injury.

Scatterplot of the Results. When all possible Primary/Secondary covariate patterns are considered, 56 individual injury groups are developed. The agreement between predicted survivability and observed outcomes (fatality/incidence) for the 56 groups is illustrated in Fig. 1. Each dot represents one of the 56 injury groups. A point that lies close to the diagonal line indicates that the model fits the data well for the group that the point represents (i.e., the predicted survivability is the same as the observed survivability). Dot sizes are proportional to incidence levels. Note that most of the dots that lie far from the diagonal are associated with lower incidence levels, which account for fewer overall cases. Tactics to obtain a better fit (better ranking techniques, more explanatory variables) are discussed later. Observed fatality/incidence ratios and computed model probabilities for each injury group are listed in the appendix in order of descending threat-to-life.

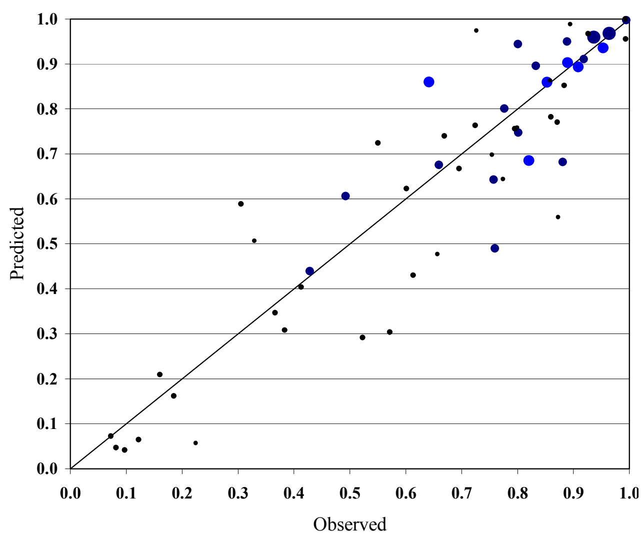

Figure 1. Observed fatality/incidence vs. predicted survival probability. Dots represent injury groups. Dot sizes are proportional to group incidence levels.

Basis of analysis: Deviance. The predictability of the Primary/Secondary model may be compared against MAIS and other survivability prediction indices by using deviance. Deviance is based on a likelihood ratio of the proposed model versus a saturated model. A saturated model is one that contains as many parameters as there are data points. (An example of a saturated model is fitting a linear regression model when there are only two data points.) In mathematical terms, deviance, D, is the comparison of observed (y) to predicted (π) values using the likelihood function (Hosmer and Lemeshow, 1989):

$$D = -2\ln\left[\frac{\text{likelihood of current model}}{\text{likelihood of saturated model}}\right]$$

$$= -2\sum_{i=1}^{n}\left[y_i \ln\left(\frac{\pi_i}{y_i}\right) + (1-y_i)\ln\left(\frac{1-\pi_i}{1-y_i}\right)\right] \quad [2]$$

The quantity inside the brackets is the likelihood ratio. Models with lower values of D have better predictability. Generally, as more variables are added to the model, its predictive ability increases. If a model were constructed using every variable in CDS, its predictive ability would approach 100%, the bracketed term would approach unity, and D would approach zero. The significance of a model improvement can be seen using the G-statistic in which the value of D given by the improved model is compared with the value of D for a baseline model:

$$G = D(\text{baseline model}) - D(\text{improved model})$$

$$= -2\ln\left[\frac{\text{likelihood of old model}}{\text{likelihood of new model}}\right] \quad [3]$$

The G statistic plays the same role in logistic regression as does the numerator of the partial F test in linear regression. To understand the role of G, consider a new, hypothetical model with one variable plus an intercept (β_0 and β_1) and a baseline model with just the intercept (β_0). Under the null hypothesis, β_1 is equal to zero, and the G statistic will follow a chi-square distribution with 1 degree of freedom. In order to reject the null hypothesis (that the β_1 coefficient is not different from zero) at the 95th percent confidence level, the p-value for the test $X^2(G, \text{DOF}=1)$ must be less than 0.05. For new models having the same DOF, a higher G-statistic means the fit is better.

Comparison with other models. Table 3 compares survivability predictability of the Primary/Secondary model with MAIS, the Injury Severity Scale (ISS), and Anatomic Profile (AP). ISS is a CDS variable that is computed for each crash victim by sorting the victim's injuries into eight body regions and finding the highest AIS severity level in each region. Of these eight levels, only the three highest are used to determine the ISS, which is computed by summing their squares

(Baker et al, 1974). AP is determined by sorting a victim's injuries into three groups: Group A – head (non-face) and spinal cord; Group B – thoracic and front of neck; Group C – all other injuries. A "Group ISS" is then computed for each group. That is, the three highest AIS levels within each group are squared and summed, resulting in three "Group ISS" values: ISS-a, ISS-b, and ISS-c (Copes et al, 1990). Also shown is the effect of adding "Age" into the Primary/Secondary model.

The models in Table 3 are ordered by increasing ability to predict survival. The improvement of each model over the baseline model is statistically significant according to the *G* statistic. The only non-baseline models whose significance of improvement may be directly comparable are the last two, since they are the only two that are nested. Their *G*-statistics indicate that there is significant improvement when adding the "Age" variable to the to Primary/Secondary according to $X^2(G,19) < 0.05$.

Table 3. Comparison of models used to predict fatalities.

Model	No. of Parameters	Deviance	G-Statistic
Baseline $\text{Logit}[\beta_o]$	1	616,429	—
Maximum AIS $\text{Logit}[\beta_o + \beta_i{*}MAIS]$; i=3,4,5	4	507,629	108,800
Injury Severity Scale (Baker et al, 1974) $\text{Logit}[\beta_o + \beta_1{*}ISS]$	2	499,058	117,371
Anatomic Profile (Copes et al, 1990) $\text{Logit}[\beta_o + \beta_1{*}ISS_a + \beta_2{*}ISS_b + \beta_3{*}ISS_b^2 + \beta_4{*}ISS_c^2]$	5	473,110	143,319
Primary/Secondary $\text{Logit}[\beta_o + \beta_i{*}P_i + \beta_j{*}S_j]$; i=1..9; j=10..19	19	470,258	146,171
Primary/Secondary, Age $\text{Logit}[\beta_o + \beta_1{*}Age + \beta_i{*}P_i + \beta_j{*}S_j]$; i=2..10; j=11..20	20	463,444	152,985

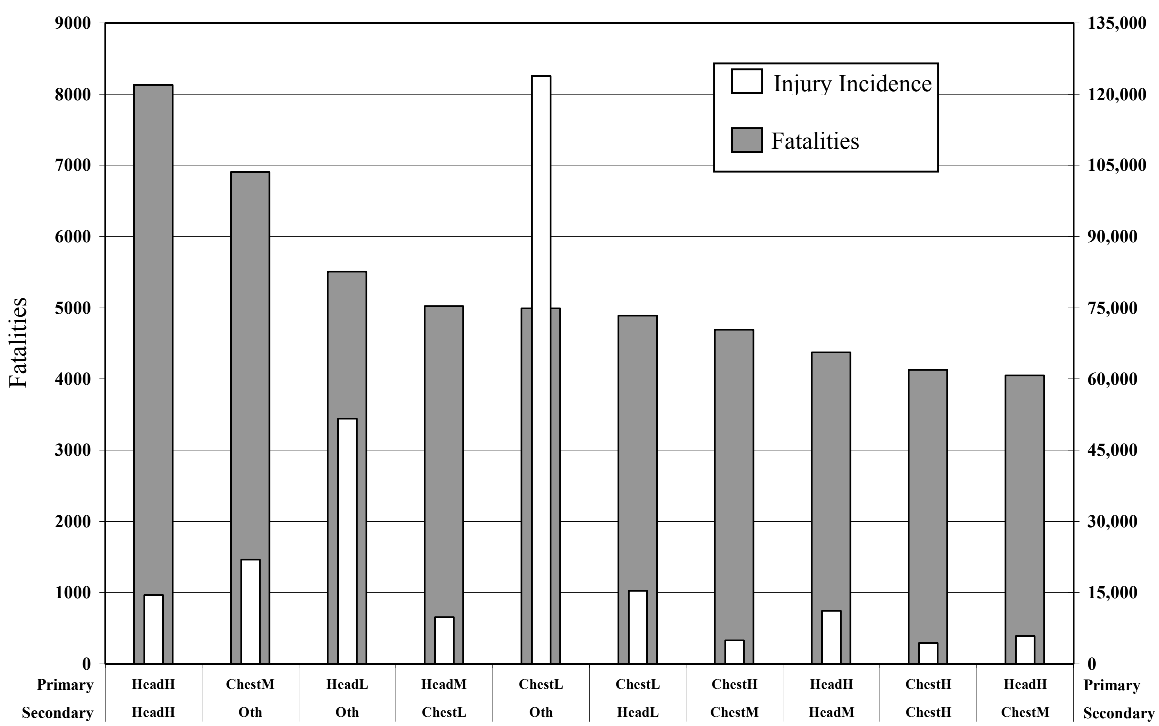

Figure 2. Injury incidence and fatality levels for the top ten injury groups in terms of fatalities.

Top ten list. Of the 56 Primary/Secondary groups, the top ten in terms of the total fatalities are shown in the Figure 2 bar chart. Head and chest injuries are predominant. These ten groups represent 50% of all fatalities, and most have relatively low incidence levels. Those having the lowest incidence-to-fatality ratios are candidates for further consideration – the working data set may be queried to obtain the six-digit codes for a more comprehensive understanding of the injury specifics.

Options: tertiary injuries. As an alternative to the Primary/Secondary model, a model having a third, fourth, or fifth injury could be used. For example, a Primary/Secondary/Tertiary model produces a slightly lower (though not significantly lower) deviance than the Primary/Secondary model alone. However, confidence in the individual parameter values diminishes considerably (the standard errors of the β_i's are greater than the values the β_i's themselves). With a tertiary injury model, there are 219 injury categories (instead of 56) to populate, and there are not enough observations in each category to obtain a reliable fit. On the other hand, if the number of tertiary injury groups is limited to two or three, then a good-fitting model may be possible.

DISCUSSION

The Primary/Secondary survivability ranking system offers a new utility over MAIS and other indices. If one desires to estimate the number of lives saved if a particular injury is mitigated, it may be accomplished forthrightly under the Primary/Secondary scheme because particular injuries can be singled out. This, however, is much harder to accomplish in the context of MAIS or ISS because individual injuries are not isolated. Figure 2 – which shows the injuries associated with the majority of fatalities – provides a template of how the new injury characterization method may be used in to prioritize research activities. It helps identify particular injuries that may be mitigated via some countermeasure, resulting in a significant number of lives saved.

Injury Discrimination. As stated earlier, efforts to prevent abdominal and lower extremity injuries are difficult to justify using MAIS reasoning because they are not normally the primary, MAIS injury. Usually, fatally injured crash victims who suffer from such injuries also sustain at least one other injury with a higher AIS ranking. However, abdominal and lower extremity injuries do affect the risk of fatality – they have a secondary role and act as "fatality modulators". Under the Primary/Secondary scheme, the contribution of such lesser injuries may be accounted for directly. Referring to Table 1, suppose all "LowEx" and "$Abdo_L$" injuries were prevented. To compute a lives-saved estimate, the incidence levels of groups involving these two injuries must be redistributed into other, lesser-severity groups. In this instance, there is just one lower-severity group – the "Oth" group – so all injuries would be redistributed into that group. Using the known survival probabilities of the re-populated groups, the number of lives saved may then be computed as shown in Table 4, where it is estimated that: 779 – 547 = 233 lives would have been saved in 2001.

Table 4. Computation of lives saved in 2001 if all LowEx and $Abdo_L$ injuries (ref: Table 1) were eliminated.

Primary Injury	Secondary Injury	Estimated Fatalities	Hypothetical Fatalities Assume $Abdo_L$, LowEx $\rightarrow$ Oth
$Chest_H$	LowEx	0	0
$Chest_M$	LowEx	271	218
ChestL	LowEx	164	127
$Head_H$	LowEx	14	11
$Head_M$	LowEx	15	12
$Head_L$	LowEx	49	39
$Abdo_H$	LowEx	2	1
$Abdo_L$	LowEx	1	0
LowEx	LowEx	103	0
$Chest_H$	$Abdo_L$	7	7
$Chest_M$	$Abdo_L$	54	48
$Chest_L$	$Abdo_L$	83	71
$Head_H$	$Abdo_L$	0	0
$Head_M$	$Abdo_L$	4	4
$Head_L$	$Abdo_L$	0	0
$Abdo_H$	$Abdo_L$	11	9
$Abdo_L$	$Abdo_L$	1	0
Totals		779	547

Model Improvement Tactics. The injury groups in Table 2 represent a trade-off between generality and specificity. Survivability models built from more descriptors usually have better predictive power (deviance is lower) because they are more saturated. Not surprisingly, AP (which uses information on nine injuries) outperforms ISS, and ISS (which uses information on three injuries) outperforms MAIS in terms of survivability prediction. On the other hand, the Primary/Secondary scheme – by virtue of its more objective ranking of injuries – outperforms all others even though it relies on just two injuries. Moreover, if "Age" is added to the Primary/Secondary model (or any of the other models), then its predictive ability becomes even greater.

The "Age" variable has a high report rate in almost all databases, including the CDS, and it may be included in the model directly. Other variables may also improve the predictive ability, but they may confound the model if there are a sizeable number of cases in which they are unknown.

The working data set, itself, may also be changed in order to obtain a better model. Two injury groups, [$Chest_L$,Oth] and [LowExt,Oth], contain about a third of all cases and both have very high survival probabilities (about 95%). The model fits these two groups very well (they are associated with the two biggest dots in Fig. 1). This means that the fit of other groups that contain the same types of injuries (and have much lower group survival probabilities and therefore strike more interest) will suffer. Despite the high survival rate, these two groups are associated with a significant number of fatalities, so omitting them outright from the data set is not recommended. Instead, it may be better to re-define the "$Chest_L$" and "Oth" variables into new subvariables. Then new injury groups can be formed and ones having high survival rates and few associated fatalities may be omitted from the working data set.

Redefined Injury Groups. As presented, the injury groups represented in the Primary/Secondary model are used to perform a general analysis that is not directed at studying any particular injury or countermeasure. Oftentimes a more specific analysis is desired. For example, consider a hypothetical side air bag designed to mitigate pelvic injuries. A modified Primary/Secondary model may be used to predict the number of lives saved by the air bag. New injury groups may be defined that are specific to the injury groups that are affected by the countermeasure (abdomen, hip/pelvis, and upper femur). Incidence levels for the new groups may be extracted from a data set that includes only side impacts to near side adults. An analysis akin to the one demonstrated in Table 4 will provide an upper limit of the number of lives that could possibly be saved by the hypothetical air bag. A closer inspection of these groups will reveal the precise nature of the pelvic injuries. This helps determine whether instrumentation in a dummy used to test the new air bags will adequately pick up the injuries.

To get a more realistic estimate of lives saved by the new hypothetical air bag, some level of injury reduction must be known (presumably from tests with dummies). Then, given a reduction in pelvic injury severity, a judgment must be made of how the various injury groups are to be re-populated. Once new incidence levels are established, an estimate of lives saved may be computed directly, as is demonstrated in Table 4.

Impending Work: Ranking of Injuries Objectively. The selection of the primary and secondary injuries in the study herein is partly based on AIS severity rankings. The assignments of the 1-6 severity suffixes to specific injury codes are provided by the AIS coding manual and are based on a panel of experts commissioned by the Association for the Advancement of Automotive Medicine. Over the years, AAAM panels have reconvened and new injury coding systems have emerged. However, the AIS severity rankings are more or less held over from the original ones that were assigned back in 1976 (AAAM, 1976).

A multi-year accumulation of CDS injury data now makes it possible to use actual mortality outcomes to objectively rank specific types of injuries by survivability. Objectively ranking injuries would help discriminate among injuries that share the same AIS severity. It would also sort out questionably ranked codes, such as those coded as "Not Further Specified" (NFS). Within the working dataset used in this study (about 7000 crash victims), about 20% of all AIS 3+ injuries make use of an NFS code. NFS codes are used when detailed medical information is lacking. For example, a medical record for a particular crash victim who has, say, a skull fracture may lack specific details such that the fracture is given a special NFS code. NFS injuries are always ranked at an AIS level that is equal to or lower than the same general injury that is described more specifically. This minimum severity rule may not always reflect the true severity of the injury.

Establishing the objective injury ranks would require a balance between objective statistics (such as incidence-to-fatality ratios derived from the actual data) and common-sense heuristics (e.g., a "moderate" laceration cannot be ranked higher than a "major" laceration). Such an objective ranking system would help identify the two injuries of each crash victim that are – on average – statistically the most life-threatening, thereby improving the predictability (deviance is decreased) of the Primary/Secondary model.

Imputation in DOA Cases. As mentioned earlier, crash victims who are "dead on arrival" (DOA) often have an incomplete injury record with only injuries of

MAIS= 2 listed. These cases were excluded from the present analysis. However, this exclusion effectively biases the working dataset in favor of non-fatal injuries. Rather than excluding DOA/MAIS=2 cases outright, a better solution may be to somehow impute a more severe primary injury. Methods to impute missing data within NHTSA's databases have been developed for certain variables. For example, Rubin et al (1998) describes a method to impute blood alcohol levels that are missing from fatality crash data. It may be possible to develop a method to impute injuries in DOA's.

Standard Errors. Unlike the example provided herein, a comprehensive analysis ought to provide confidence intervals for the number of lives saved or injuries prevented. For this reason, it may be helpful to combine a CDS model with incidence levels obtained from the General Estimates System (GES), a more general epidemiological crash database maintained by NHTSA. GES estimates are more trustworthy because they have much lower standard errors by virtue of higher sampling rates. Moreover, properly computed standard errors that take into account the CDS (and GES) multi-stage sample may be found with the SUDAAN LOGISTIC procedure (Shah, Barnwell, and Bieler, 1996) using the sampling weights and sample stratifiers in the data set.

SUMMARY

A procedure is presented to estimate the risk to life that multiple injuries pose to crash victims. The procedure is intended to be used in an aggregate study of CDS crash victims. The new procedure uses only the two most serious injuries to characterize a victim's entire injury record. This new Primary/Secondary model has an improved predictive ability over MAIS and ISS partly due to an objective scheme that ranks individual injuries using actual crash outcomes.

The biggest advantage the Primary/Secondary model, however, is its ability to discriminate among several types of injuries. A general injury model – with 56 different injury characterizations – is demonstrated herein. For a given primary injury, the threat-to-life is shown to vary profoundly depending upon the type of secondary injury. Moreover, a 56-injury group breakdown provides a means to attribute fatalities to specific injuries, including those that are relatively non-threatening such as lower extremity injuries.

The survival predictability of the Primary/Secondary model may be improved substantially by including "Age" as an additional variable. The model would also benefit from a more thorough and objective scheme to rank injury severity. Furthermore, the model may be adapted for an assessment aimed at a particular countermeasure or injury.

REFERENCES

Association for the Advancement of Automotive Medicine (1990). The Abbreviated Injury Scale, 1990 Revision. Des Plaines IL: Association for the Advancement of Automotive Medicine.

American Association for Automotive Medicine (1976). The Abbreviated Injury Scale, 1976 Revision Arlington Heights IL: American Association for Automotive Medicine.

Baker SP, O'Neill B, Haddon W, Long WB (1974), The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care, J Trauma V14(3):187-96.

Benton, EJ (1993). NASS 1993 Injury Coding Manual. Washington DC: Dept. of Transportation (US), National Highway Traffic Safety Administration. Report No.: DOT HS 807 969; 1993.

Blincoe L, Seay A, Zaloshnja E, Miller T, Romano E, Luchter S, Spicer R (2002). The economic impact of motor vehicle crashes, 2000. Washington DC: Dept. of Transportation (US), National Highway Traffic Safety Administration. Report No.: DOT HS 809 446.

Copes WS, Champion HR, Sacco WJ, Lawnick MM, Gann DS, Gennarelli T, MacKenzie E, Schwaitzberg S (1990), Progress in characterizing anatomic injury, J Trauma V30(10): 1200-1207.

Eppinger R (1987), On the examination of biases in the Abbreviated Injury Scale, Proc. 11th International Conference on Experimental Safety Vehicles, Washington DC, May 1987.

Eppinger R, Partyka SC (1981), Estimating fatality reductions associated with safety improvements, Proc. 8th International Conference on Experimental Safety Vehicles., DOT HS 805 555, October 1981.

Hosmer DW, Lemeshow S (1989), <u>Applied Logistic Regression</u>, John Wiley and Sons, New York.

National Highway Traffic Safety Administration (2002), Preliminary regulatory analysis: Proposed amendment to FMVSS No. 213 frontal test procedure, Washington DC: Dept. of Transportation (US), National Highway Traffic Safety Administration.

Rubin DB, Shafer JL, Subramanian R (1998). Multiple imputation of missing blood alcohol concentration (BAC) values in FARS, Washington, DC: U.S. Department of Transportation. Report No. DOT HS 808 816.

SAS Institute Inc. (1999), SAS/STAT User's Guide, Version 8, Cary, NC: SAS Institute Inc.

Sartorelli KH, Rogers FB, Osler TM, Shackford SR, Cohen M, Vane DW (1999), Financial aspects of providing trauma care at the extremes of life, J Trauma 1999 Mar;46(3):483-7.

Shah BV, Barnwell BG and Bieler GS, (1996). SUDAAN User's Manual: Release 7.0, Research Triangle Park, NC: Research Triangle Institute.

Siegel JH, Gens DR, Mamantov T, Geisler FH, Goodarzi S, MacKenzie EJ (1991), Effect of associated injuries and blood volume replacement on death, rehabilitation needs, and disability in blunt traumatic brain injury, Crit Care Med 1991 Oct;19(10):1252-65.

APPENDIX

Injury Groups – Survivability probability and 1993-2001 incidence level (weighted).

Injury Group (Primary,Secondary)	Model Prob. Logit[$\beta_o + \beta_i * P_i + \beta_j * S_j$]	Observed Incidence/ Fatalities	Observed Incidence (weighted)
ChestH, ChestH	0.072	0.072	4451
ChestH, ChestM	0.082	0.047	4922
ChestH, HeadH	0.097	0.042	2840
ChestH, HeadM	0.122	0.065	1742
ChestH, AbdoH	0.160	0.209	1308
ChestH, ChestL	0.185	0.162	2505
ChestH, HeadL	0.224	0.057	578
ChestH, LowEx	0.305	0.589	1835
ChestH, AbdoL	0.329	0.507	243
ChestH, Oth	0.366	0.347	2448
HeadH, ChestM	0.383	0.308	5857
ChestM, ChestM	0.413	0.404	5367
HeadH, HeadH	0.428	0.439	14496
HeadH, HeadM	0.492	0.606	11116
ChestM, HeadM	0.523	0.291	5444
AbdoH, ChestM	0.550	0.724	3763
HeadH, AbdoH	0.571	0.304	1028
ChestM, AbdoH	0.601	0.623	1420
HeadH, ChestL	0.613	0.430	3583
ChestM, ChestL	0.642	0.860	19295
AbdoH, HeadM	0.656	0.477	363
HeadM, HeadM	0.659	0.675	9535
HeadH, HeadL	0.669	0.740	4581
ChestM, HeadL	0.695	0.667	3578
AbdoH, AbdoH	0.724	0.764	1833
HeadM, AbdoH	0.726	0.974	434
HeadH, LowEx	0.754	0.698	658
AbdoH, ChestL	0.757	0.643	8551
HeadM, ChestL	0.759	0.490	9852

Continued.

APPENDIX, continued.

Injury Groups – Survivability probability and 1993-2001 incidence level (weighted).

Injury Group (Primary,Secondary)	Model Prob. Logit[$\beta_o + \beta_i * P_i + \beta_j * S_j$]	Observed Incidence/ Fatalities	Observed Incidence (weighted)
HeadH, AbdoL	0.774	0.644	56
ChestM, LowEx	0.776	0.801	7773
ChestM, AbdoL	0.795	0.756	1743
AbdoH, HeadL	0.799	0.758	225
HeadM, HeadL	0.801	0.944	14525
HeadH, Oth	0.801	0.747	6204
ChestM, Oth	0.820	0.685	21937
HeadL, HeadL	0.833	0.896	13601
ChestL, ChestL	0.853	0.859	22109
AbdoH, LowEx	0.858	0.863	450
HeadM, LowEx	0.859	0.782	916
AbdoH, AbdoL	0.871	0.770	993
HeadM, AbdoL	0.872	0.559	196
ChestL, HeadL	0.881	0.682	15370
HeadL, LowEx	0.883	0.852	3183
AbdoH, Oth	0.888	0.950	6945
HeadM, Oth	0.889	0.903	31706
HeadL, AbdoL	0.894	0.988	338
HeadL, Oth	0.909	0.893	51629
ChestL, LowEx	0.918	0.911	16948
ChestL, AbdoL	0.926	0.967	5160
ChestL, Oth	0.937	0.960	123865
LowEx, LowEx	0.953	0.936	22663
LowEx, Oth	0.964	0.968	99945
AbdoL, LowEx	0.993	1.000	1785
AbdoL, AbdoL	0.993	0.956	1247
AbdoL, Oth	0.994	0.997	11768